对人工智能大模型能力边界的初探和一种新的AGI实现途径

曾婷[1*]，陈永聪[2]，陈星月[3]
(1，清华大学图书馆，北京，100084，E-mail: ceng-t@mail.tsinghua.edu.cn；2，新千年未来科技（北京）有限公司，北京，100084；3,清华大学附属中学，北京，100084)

摘 要：

目前主流的人工智能，普遍采用“注意力机制 + 深度学习”+“强化学习”的技术道路。我们认为“强化学习”无法适用到那些难以大量“试错”的领域。所以，要想实现能适用于任何领域的通用人工智能，我们必须转变实现道路。所以，我们提出了一套不同于“深度学习+强化学习”的机器学习方案，它通过小样本、累积学习，同样实现了和transformer相似的注意力机制，也同样创建了全连接知识网络。并且，它不需要采用试错学习的方式，就可以实现和环境的互动决策。并且人类可以给它预置不同的先天需求，来实现多目标平衡，从而实现远高于目前人工智能的安全性。在本文中，我们提出了一套从0到1的新机器学习技术方案。

关键词： 大模型 ChatGPT GPT-4 通用人工智能 AGI

1，引言

目前主流的人工智能，普遍采用“注意力机制 + 深度学习”+“强化学习”的技术道路。其中“注意力机制 + 深度学习”主要用于建立知识网络；而“强化学习”主要用于提高机器和环境的连续互动决策能力。比如Google公司2022年6月推出的“GaTo”，单个模型可以完成600多种不同的任务[1]。再比如Open-AI推出的大模型GPT-4，在知识理解和推理方面，获得了惊人的进步[2]。但目前人工智能大模型的能力上限在哪里？“注意力机制 + 深度学习”+“强化学习”能实现真正的“通用人工智能”吗？我们在本文中进行了一些初步探讨。

2，目前大模型的能力上限

首先，我们注意到目前“强化学习”都是采用外部反馈来告诉机器在不同决策路径下结果的“好”和“坏”，这种外部反馈可以来自预设的奖励函数，比如alpha go中，反馈来自于判别输赢的奖励函数；也可以是人类的反馈，比如目前大模型中普遍采用的RLHF（Reinforcement Learning from Human Feedback）技术[3]。所以，目前“强化学习”的本质是先“尝试”、后“反馈”的方式，让机器通过不同的决策路径，获得不同的“奖励”信息。然后，机器才可以获得“自身状态”和“不同策略”之间的利弊关系，从而获得和环境的互动决策能力。[4]

我们认为，当机器面对需要和环境互动决策时，目前“强化学习”这种方式，更加像一种“进化”学习的方式，它的本质是“试错”，通过外部反馈来实行淘汰。这和生物的进化非常相似。所以，这样的学习方式只适用于虚拟环境中，比如游戏、元宇宙和内容生成任务中。对于必须在实际环境中进行互动学习的任务，

比如照顾老人，驾驶车辆，强化学习难以适用。

其次，“深度学习”所创造的知识，人类难以理解。这会导致机器和人类是两套知识体系，彼此无法理解！比如，人类难以理解大模型的决策过程。同样，大模型也难以真正地理解人类语言所代表的知识。比如，目前还没有任何人工智能，可以读一遍“面包机”的使用说明，就能在不同的面包店，操作面包机烤面包[5][6][7]。而人类却可以通过阅读“面包机”的使用说明，就获得他人已经积累的关于面包机的使用知识。从而在面对“操作面包机”这个新任务时，人类可以在这些已有决策知识指导下进行决策，和环境互动，比如，人类在打开面包机顶盖时，不需要先去尝试各种方案（比如砸开面包机的顶盖），而是通过语言直接获得人类已经积累的经验。

所以，我们认为，真正的机器学习，应该像人类一样，面对一个新任务，可以根据自己过去的经验，来预测不同决策路径下的“好坏”，最多选有限几个方案去尝试，就可以获得处理新任务的决策知识。更进一步，我们认为真正的学习，也应该和孩子学习方式类似，通过语言来直接获得人类已经积累的经验。在面对新任务时，一次尝试都不需要，直接一次成功！比如在实验室里，老师教孩子们做实验时，是通过语言传授，直接把人类已有的决策经验传递给孩子们。孩子们可以在获得老师传递过来的知识后，在不同的环境下，通过一步步决策，和环境互动，就可以直接完成实验。尽管孩子们可能是第一次做这些实验！

所以，我们认为，因为知识结构和学习方法不对，导致目前人工智能大模型无法解决下面的严重缺陷：

2.1，不能自主解决问题。

比如目前人工智能，它看到主人摔倒时，并不会主动过来帮忙[7]。

这是因为机器没有自己的需求，就不可能产生自己的目标。由于机器没有自己的目标，就不可能主动创建一个任务。所以，面对实际的社会生活中层出不穷的意外情况，机器只能通过预设流程来处理。它的预设流程，可以来自外部人类编写的 prompt 流程[8]。也可以是机器按照预设指令，使用输入信息在知识库中寻找相似流程来整体模仿。

目前一些人也在试图编写递归调用函数。比如在一个任务下，调用“寻找过去完成相似任务的流程”的函数，然后针对获得的每一个流程，再次调用“寻找过去完成相似任务的流程”的函数，来获得更多的流程。这是一种递归调用的方式。目前具有代表性的项目是 AUTO GPT，它通过存储把从搜索引擎获得的知识，和从 GPT-4 获得的知识，聚合起来，然后作为 prompt，从 GPT-4 寻找解决方案。通过这个过程的递归调用，就可以完成一个相对宏观的任务。如果任务需要的所有流程和参数都已经存在于大模型的函数库中，大模型可以完成任务。但 AUTO GPT 能完成的任务，并不会超越大模型能完成的任务的边界。而且，在现实生活中，层出不穷的意外情况，很可能会导致在某一个点上，并没有合适的流程和参数匹配目前的环境，这就会导致递归调用失效[9][10][11]。

所以，大模型不会自发地去创建新流程！大模型本质就是一种高级的编程语言。使用的语法就是自然语言。如果我们理解了这一点，就能看清楚大模型的边界在哪里。无论我们添加多少高级函数到大模型中去，也就是无论我们集成多少工具、APP 到大模型中去，大模型都不会自发的去创建新的流程。它所有的流程都是模仿过去的流程（递归调用），或者人类为其预置的流程（主程序调用）。这两种方式，本质上都是“使用预置的流程来处理所有问题”。无论这个流程中

有多少 if...else...，考虑到多少种可能性，它都是预置的，预先就存在的。它不是面对具体任务，临时创造出来的！

所以，大模型的本质就是给人类提供一套编程平台，使用的编程语言就是人类语言。而之前的编程语言是 C++，Jave，Python 等计算机语言。同时，大模型也提供了一套函数，这套函数可以使用人类语言来调用。比如"创建 PPT"函数，就是人类语言命令 + 该函数需要的各种参数"。大模型在命令 + 参数的模式下，完成任务。这样，未来人人都可以是程序员。而对大模型更加了解的人，可以编写更好的程序，这些人就被叫做"Prompt engineer"，他们是未来的程序员。

目前大模型还在不断的丰富自己的函数。比如未来，它的函数输入参数还可能包含图像、视频、动作，或者任何传感器输入，函数的输出也可能包含图像、视频、动作，或者任何其他多模态序列[13]。所以未来，所有流程都可以采用大模型的编程语言（自然语言）来编程。大模型就是一个机器语言平台，它和 Python 没有本质区别。

比如，在 Python 程序中，人类把思想变成 Python 程序流程，python 平台负责把函数翻译成各种下层驱动能够理解的机器程序，来完成任务，从而实现人类预设的，解决问题的流程。

所以大模型就是新一代的 Python。Python 比 C++更适合人类接口，所以它更流行。GPT-4 比 Python 更适合人类接口，所以它未来会更流行。但 C++依然没有消失，因为在特定领域，它有它的优势。所以，未来各种专用的大模型，或者 Python，也不会消失，它们还有少量存在的空间。所以，未来所有的软件都可以用大模型语言来重写，而懂得大模型编程的新一代程序员（Prompt engineer）将会成为新一代的"白领"[14]。

所以，大模型的核心本质，就是更加友好的一套"编程语言"。它接口是自然语言，函数丰富多样，而且是多模态的，预期未来很快就会建立起自己的生态圈。但它的本质，依然脱离不了编程语言！它是入门门槛极低，功能极大扩张的一套编程语言，这是它的本质。所以，未来以大模型为基础，可以完成任何可以预测流程的任务。那些复杂的任务，无非就是多编写一些 if...else...分支。但现实生活中，存在大量无法预测流程的任务！比如照顾老人、开车做饭、陪伴孩子、农业生成等等。人类可以给帮忙考虑到各种可能性，但在和真实环境互动的过程中，总有意外情况出现。这时，机器怎么处理这些意外情况，人类就无法预测。所以这会非常危险，尤其是在大模型的函数涉及到人类生活的方方面面时，可能给人类带来无法接受的损失。

2.2，知识无法实时更新。

目前人工智能，采用大数据训练，知识无法实现实时更新。而知识的实时更新，对于和环境互动的机器而言，至关重要。因为机器和环境的互动，就是机器获得新知识的过程。如果机器所获得的知识无法实时更新，就会导致机器无法实时根据环境的反馈来更新自己的决策知识。所以，这样的机器，面对同样的环境，就会不断犯同样的错误[3]。

2.3，无法适用于需要和真实环境互动的领域。

在需要和真实环境互动的领域，比如自动驾驶、做家务、护理病人等领域，机器需要建立自己行为和外界环境之间的互动决策知识。而这些领域无法大量的试错，所以机器无法通过强化学习，在真实的环境中通过互动来建立这些领域的

决策知识。所以，目前人工智能的技术方案，无法适用于这些领域[2]。

3，目前大模型的能力初探

3.1，如何描述一个矩阵包含的信息的？

如何来描述一个矩阵所包含的信息的？虽然一个矩阵可能包含很多矢量，但我们首先关心的是：有多少个独立的矢量？也就是说，矩阵的秩是多少？然后，我们可以通过矩阵对角化，找到对应的一套特征向量，它们数量上就是矩阵的秩，它们也是这个矩阵的一套坐标基底簇。这套坐标基底簇是完备的，是正交的，是这个矩阵的最简洁描述。矩阵中任意矢量，都可以通过这套坐标基底簇来表达。

如果我们建立的坐标基底簇并非正交的，但如果是完备的，那么我们同样可以用这套坐标基底簇来表达矩阵中任意信息。如果坐标基底簇是非完备的，那么矩阵中就存在一些矢量，无法通过这套坐标基底簇来表达，这时就需要增加坐标基底簇的维度。

那么，我们是如何识别矩阵中一个矢量包含那些信息呢？显然，如果是完备基底簇，每个基底就是一个维度，任意矢量向坐标基底投影，就能准确的获得一个矢量所包含的所有信息。如果基底坐标簇是本征正交基底，那么我们实现了用最简洁的系数来表达这个矢量的全部信息。如果基底坐标簇不是完全正交的，那么我们希望它尽可能接近正交基底簇，因为这时我们获得的系数矩阵是稀疏的。通过稀疏化的系数矩阵，配合一套基底坐标簇，我们就能理解矩阵中的任意矢量包含的信息了，而且这时信息分量之间相对独立。所以两个矢量之间的关系，就体现在它们坐标基底分量之间的关系。

所以，如果在一个矩阵中，两个矢量存在非零的同一个维度上的分量，我们则认为它们之间存在局部相似性。我们可以认为这两个矢量之间存在某种连接关系。如果两个矢量，存在多个局部相似性，我们可以认为这两个矢量之间存在更强的某种连接关系。如果类似的某种连接关系在矩阵中反复出现，我们就可以认为这种连接是一种普遍规律。这就是知识。

所以，在一个矩阵中，寻找知识，就是寻找其中所有的普遍规律。规律本身是由矩阵中部分基底分量来表达的，既包括维度，也包括维度上的大小。它是从大量的矢量彼此连接关系中，获取的共有连接关系。所以它包含的维度更低，代表的范围更广。所以知识是从大量的矢量关系中概括出来的。每一个知识本身，也是矩阵中一个矢量。它们可以用一个系数矩阵来表达。而大量的这种知识，就构成了知识网络。

所以，如果我们需要发现一个矩阵的全部知识，那么，最重要的就是：找到一套基底坐标簇，然后用它来分解矩阵中任意矢量；然后通过系数矩阵之间的关系，来获得不同矢量之间的连接关系；然后找到那些能重复出现的连接关系；它们就是从矩阵信息中提取出来的知识。

但一个矩阵可以有多套基底坐标簇。只要维度不低于矩阵的秩，任意一个基底坐标簇本质上都可以采用。为了获得最简洁的知识体系，我们采用正交坐标簇就可以了。但如果矩阵非常庞大，找到一套正交基底簇的计算量难以实现。这时，我们就可以沿那些重复性最高的知识，用它们作为基底坐标簇。这样至少矩阵中大部分知识可以采用稀疏矩阵来表达。也就是说，我们获得了一套可以简洁表达矩阵中常见知识的坐标基底簇。

有了基底坐标簇，任意矢量都可以分解到基底坐标簇，采用系数矩阵来表示。

任意矢量之间的相似度，都可以通过它们的空间距离来表示。而空间距离可以采用类似欧式空间距离来表达。任意矢量之间的映射关系，都可以通过系数矩阵的映射关系来实现。

3.2 深度学习是如何创建知识的？

假设在一个外星世界里，有一群人，建立了这样一个 4 维的信息空间，它包含大量的图像、声音和动作。在那里的人类，她们是通过像素、音节和动作模式来作为感知能力底线的。所以，面对纷繁复杂的像素、音节和动作模式的各种组合，她们无法看明白彼此之间的连接关系。

于是，她们就开始采用试错法，尝试不同的基底坐标簇，希望找到一个特定基底坐标簇，在这样的基底坐标簇下，她们感兴趣的像素、音节和动作模式组合，可以形成彼此分离的一簇一簇的。那么，这个基底坐标簇就是她们需要的。

由于原始数据的维度数量非常高，比如 64*64 图像，它们的维度就是 64*64，它们就是 64*64 个二维脉冲函数作为基底的原始基底坐标簇。而她们的目标维度则可能要小的多，可能远小于 64*64，这是因为她们的目地是找到常用特征组合，而不是需要表达所有信息，必然有一部分信息会损失，她们不可能把所有原始信息都放到最终的维度上去。所以，一种可能的试错法，就是每一次尝试一种坐标基底簇后，就把变换后的部分维度信息丢弃或者压缩。然后和目标进行对比，看看误差是增加了还是减小了，然后决定下一次改变坐标基底簇的方向。显然，每尝试一次，信息就会损失一部分。经过多次尝试后，信息的损失过大，可能会导致其中有用信息的损失，从而无法完成任务。所以，有用信息在整体信息中所占比重，每一次变换后信息损失率，就决定了最多能变换几次。而在有限的变换尝试中，机器可能难以找到最优的坐标基底簇。所以，一种可行的解决方案就是每一次变换后，把损失的信息一部分补回来，这样就能增加变换层数，从而提升找到最优的坐标基底簇的概率。这就是残差网络。当然，也可以考虑在多次神经网络的映射过程中，插入弱非线性函数，来增加可以变换的次数。

在通过试错法寻找最优的坐标基底簇中，寻找的方向是误差的缩小，实现的工具是 BP 算法。出现在神经元中的数据实际上是基底坐标簇下的系数矩阵，而基底坐标簇本身是隐含的，没有出现在多层神经网络中。层间的变换系数，是从一个隐含基底到另外一个隐含基底变换中的坐标系数变换矩阵。BP 算法就是通过调整坐标系数变换矩阵，来从一个隐含坐标基底，走向另外一个等待尝试的隐含坐标基底。

当然，如果尝试中，选择的基底坐标簇本身非正交坐标系。可能会出现，修改一个维度上的系数时，会影响到另外一个维度上的系数。这就可能出现一种情况，那就是所有系数的调整，整体误差都不再缩小。核心原因是基底坐标簇之间非正交系带来的。如果是正交系，那么就不可能出现这种情况。所以，在尝试过程中，需要尽量走向正交坐标簇。而接近正交坐标簇的一个标志就是系数矩阵的稀疏化。所以，整个尝试的方向，需要增加系数矩阵的稀疏化约束，这就是各种正则化方法。

另外需要指出，无论是卷积，还是池化，或者其他深度学习的变种，它们的本质都没有变化。比如，卷积本质上就是一层神经网络映射，只不过在坐标系数变换矩阵中，人为地预设了大量的 0 系数。而池化，无非就是一种去掉部分维度，并采用特定非线性函数的一层神经网络映射。这就是深度学习的本质，也是在矩阵中寻找一套坐标基底的方法。

如果，空间中的数据，存在标签，那么标签的数量，就是最终需要的基底坐标簇的维度。她们寻找最终基底坐标簇的目的，就是把每一类有标签数据中，共有的最小分辨率特征（这里就是像素、音节和动作模式）的组合方式，作为每一类标签的代表，也作为一个基底坐标簇。这样获得的系数矩阵就是稀疏的。所以，我们看到，深度学习的本质，同样是在信息矩阵中，寻找一套合适的基底坐标簇。如果把需要的信息看着一个信息子空间，那么深度学习获得的基底坐标簇，就可以表达这个子空间中的任意矢量，这就是有监督学习。如果这个子空间很大，直接就包含了所有的信息，那么深度学习获得的基底坐标簇，就可以表达整个信息矩阵中的任意矢量，这就是无监督学习。如果学习的主要目的是聚类，而丢弃那些无法聚类的信息，那么这也是无监督学习。

3.3 注意力机制的本质是什么？

注意力机制的核心，就是发现信息矩阵元素之间的常见排列方式。而常见排列方式，就可以选择为信息空间的“框架”。所谓“框架”，就是它们普遍存在于信息矩阵中，采用它们作为坐标基底簇，可以简洁的描述矩阵中常见矢量。

更通俗的说，我们可以认为语言信息空间中，每一个字符都是一个维度。如果我们采用这样的坐标基底簇，我们就可以用它描述语言信息空间中任意矢量。但这样的坐标基底簇，有可能并不是最优的。如果我们把语言信息空间作为一个矩阵，那么最优的坐标基底簇，显然是由这个矩阵的特征矢量组成的基底簇。特征矢量组成的基底簇 rank 最小，描述信息最简洁。

比如，我们可以把“我今天要去参加朋友的婚礼”这句话，按照每个字符作为一个维度来分解，获得的系数矩阵就是 12 维度的。它们就是“我，今，天，要，去，参，加，朋，友，的，婚，礼”。但我们也可以把“主语...谓语...宾语”作为一个坐标基底，这个坐标基底的系数就是“我...参加...婚礼”，然后把“副词+谓语”作为一个坐标基底，这个坐标基底的系数就是“今天要去参加”，然后把“定语+宾语”作为一个坐标基底，这个坐标基底的系数就是“朋友的婚礼”，显然，后者采用了更加简洁的基底坐标簇，表达了同样的信息。而且，后者的基底坐标簇是框架性的。这些框架性的坐标簇，在不同的坐标分量下，就可以构成大量类似的信息。

而注意力机制的核心，就是建立“框架性”的坐标基底簇[16]。采用的方法，就是提取信息矩阵元素之间的常见底层框架。这个过程非常类似于人类的学习。我们学习一本书中的信息时，“先读薄，再读厚”就是同样的方法。“先读薄”就是总结出其中的框架性信息，这是一个信息压缩过程；然后“再读厚”就是在框架性信息的基础上，添加不同的细节，来构成我们创造出来的新知识，这就是一个信息创造过程[17][18][19]。

所以 transformer 类大模型，核心就是通过注意力机制[15]。注意力机制的核心目的就是获得信息矩阵中元素的常见排列方式，并按照常见程度给它们加权。越常见的排列，权重越高。那些高权重的排列方式，就是信息矩阵中，所有信息组织方式的主框架。

这个过程非常类似于通信中的信号处理过程。在时域中，看似纷乱复杂的信息，转换到频率后，它们的低频成分就决定了这个信号的大趋势，也是这个信号的主要分量。这些低频成分，就是在这一类信号中常见的组织形式。如果把每一个低频成分作为一个基底坐标簇分量，那么它们就类似于注意力机制。低频成分表达了信息之间的常见连接关系，它们是信息的“框架”。所以注意力机制，正

是通过寻找信息之间的连接方式的权重，获得了信息组织方式的“框架”。这些“框架”就是泛化的基础。“框架”和“框架”之间的映射关系，代表了“矢量”到下一个“矢量”的算法，输入“框架”+不同的细节，就是具体的输入矢量，通过“矢量”到下一个“矢量”的算法，就可以获得输出矢量，这就是知识泛化过程。

事实上，人类在学习过程中，采用了同样的方式。那些常见特征组合，就是具体“概念”，它们是权重较高的组合。它们就是空间和时间上的常见组合。从具体“概念”中进一步总结出来的常见特征组合，它们就是“抽象概念”。这个过程是可以迭代的。所以人类社会存在大量层次化的“概念”，它们就是框架。所以，抽象框架“猫”是一种常见的矩阵元素在空间和时间上的排列方式，这个排列方式中可能包含“猫”的语言、文字、声音、图像、动作、触觉等等多模态矩阵信息元素。这个排列方式中，部分矩阵元素可能拥有更高的权重，因为它们更加常见，它们可能都是“动物”这个概念。“动物”包含的元素更少，其限定范围就更小，适用范围就更大，所以在“猫”和“狗”之间，和它们共有的特征组合（比如“动物”这个概念）关联的知识就可以直接泛化。

所以，注意力机制的核心能力，就是把语言之间的统计关联性，实例化到具体的输入信息中。语言之间的统计关联性，是通过预训练获得的统计关联，这种统计关联是一种非完全统计关联。它并没有统计所有语言的排列方式中，任意语言排列组合的元素之间的关联性，因为这是不可能完成的任务。所以，在具体的排列组合下，语言之间的实际关联性，还需要基于统计关联来进一步优化。而这一步骤，就是由注意力机制来完成的。

注意力机制的核心目的，就是从统计关联性出发，采用试错法，以人类语言的前后关系作为自监督，来找到输入信息彼此，或者输入、输出信息之间的关联性，并通过权重来表达这种关联性。而这种关联性，非常类似于人类的学习总结过程。所以，机器通过深度学习（本质就是试错法），来找到一套优化的坐标基底簇。这套坐标基底簇，很可能非常靠近人类的常用概念。所以，深度学习的核心是用试错法找基底坐标簇，而注意力机制的核心是用试错法把基底坐标簇向人类概念靠近。

3.4 大模型为什么会有能力涌现？在什么时候涌现？

为什么大模型会有“涌现”现象？很简单的道理，比如一个美国人来到中国，他可以通过我们人类之间大量的共有背景信息（比如人身需求、社交结构等），通过中等数量的中英文对比，就能完成正确的翻译过程。而大模型就像一个外星人，它和人类之间并没有共同的背景信息，它看到的东西，只有人类信息之间的连接方式。所以它需要提取人类信息之间的连接方式，来预测信息的发展过程。一开始，样本不够时，它提取的“信息框架”和人类“信息框架”差异很大，所以它会不断犯错误，在黑暗中摸索，总是四处碰壁。随着样本数量的不断增加，它的“信息框架”和人类“信息框架”有更高的概率对齐。但这不是一个线性过程。比如在提升到某一个阈值之前，它就像人类语言学家解密古代语言一样，在黑暗中摸索，进展甚微。在某一个节点上，如果正确率达到阈值，整个解密过程就会大大加快，急速完成。这就是“涌现”现象。

机器“涌现”的并非智力，而是找到了“正确的信息之间常见组合方式”。这个正确的信息之间常见组合方式就是类似于人类使用的方式。因为评价它的一切标准都是人类使用的标准，所以当它的基底足够多，并且和人类基底靠近时，

它的能力就涌现出来了。

大模型的能力“涌现”，核心是通过注意力机制让大模型产生的概念和人类概念靠近了。所以训练数据必须足够，“涌现”才可能发生。因为概念靠近人类概念，所以出现了泛化能力。因为人类概念中，存在大量的抽象概念，它们是信息映射的框架。比如“猫”就是一个抽象的概念，因为它并不具体代表哪一只猫。所以，比如给机器大量的输入信息和输出信息，机器就能建立这些输入信息中的框架信息，和输出信息中的框架信息。并建立从输入框架到输出框架的映射过程。所以，输入框架+细节，就可以通过同样的映射过程，获得输出框架+细节。这就是知识泛化过程。

当训练数据量足够大，机器就可能从中发现常见复杂组合模式。常见组合模式在空间上的分布就是事物，常见组合模式在时间上的分布就是流程。常见组合模式在空间和时间上的分布，就是知识。所以大模型看上去就有了关于事物、流程的知识。

这些框架性知识，就是“世界模型”。人类正是在自己的框架性知识基础上，来认识万物，并和万物互动的。我们可以认为大模型类似于从频域来看问题。类似于面对一张图片时，我们可以采用少量的低频分量，来获得这个图片的主要内容。这就是图像压缩技术的核心。而注意力机制，其核心是通过类似的方式，采用少量的分量，来获得我们这个世界信息的主要内容。一张图片，在其低频分量上，通过配置不同的高频成分，就可以获得不同的风格调整。所以，泛化的核心就是通过“框架”配置不同的细节。

目前大模型的核心能力，是通过深度学习来建立“输入矢量”到下一个“矢量”之间的转换矩阵。有了这个转换矩阵，有了框架信息，机器就能实现知识泛化。所以，人类只要给它类似的“输入矢量”到下一个“矢量”的知识，它就可以通过模仿“输入框架”到下一个“框架”之间的转换，嵌入不同的细节来做内容生成。比如建立了“公司和创始人”这样的注意力机制，它就可以把“乔布斯和苹果”泛化到“雷军和小米”上。所以大模型是通过模仿和创造来执行任务的，这一点和人类非常相似。所以大模型可以实现小样本或者零样板来泛化，这一点都不奇怪。

所以，深度学习是一条简洁、优雅的道路，而注意力机制则是这条道路上的一处指示标，它指引人类走向智力殿堂的正确方向。而“世界模型”则是人类在这段跋涉中收获的果实。

3.5 RLHF 能最终解决大模型面临的问题吗？

目前大模型存在两个严重问题：

（一）有害内容的问题[20]。

机器的知识，人类很难理解其含义，但机器可以使用，看上去问题不大。但实际上问题很严重。最核心的问题就是：人类无法模仿机器建立的知识网络形式，给机器预置一些先天知识！这是问题的核心。因为无法给机器预置先天知识，所以不可能模仿机器建立的知识网络形式，给机器预置一些基础需求知识。机器没有自身的需求，机器就不可能有自我感知的奖励和惩罚。机器没有自我感知的奖励和惩罚，就不可能自发创建各种事物（即各种基底坐标簇组合）到自我奖励或者惩罚的投影。也就是说，机器所创建的基底坐标簇中，缺乏了奖励、惩罚、快乐、悲伤等人类有的，也必须要有的基础维度！

这是因为在基底坐标簇中缺失了这些维度，所以机器不可能把输入信息投影

到这些维度，识别出输入信息所包含的这些信息。也不可能在自己准备组合基底坐标簇作为输出时，通过使用不同的组合（也就是说机器的不同决策路径）向这些基底坐标簇投影，从而预测这些输出潜在的奖励或者惩罚。

目前大模型采用的补救方法是 RLHF。这相当于人类事后对一些训练矢量上，增加一个奖励维度的分量。也就是说，机器的基底坐标簇中，增加了一个奖励维度。如果在训练数据中，在大量不同类型，足够数量矢量上，增加在奖励维度上的分量值，就相当于建立了这些训练矢量中的共有的分量组合，到奖励维度的投影。这就是机器的奖励函数。所以，机器也可以预测不同决策下，也就是按照不同的组合方式产生的输出矢量中，包含的奖励分量。所以，机器会优选奖励分量高的输出。这就是 RLHF 学习带来的惊人效果。因为通过 RLHF 学习的知识，实际上是可以泛化的。当一个机器有了自身的奖励、惩罚维度，就有了自己初步的"趋利避害意识"，这就是为什么我们会从目前大模型看到"意识"的朦胧影子。

但这是一种事后打补丁的方式，意味做需要机器先尝试，然后人类打分反馈，它只能用于可以大量试错的领域。这类似于一个孩子博士毕业了，但完全没有"是非"观念，父母只能跟在屁股后，喊"No"，"No"，"Yes"来赋予他"是非"观念，而且他和父母还无法直接交流，只能通过"Yes"和"No"来沟通。所以，这样的学习效果效率低，而且永远可能碰到那些意想不到的 corner case！

（二）一本正经地胡说八道的问题[20]。

注意力机制是通过寻找信息之间的连接方式的权重。机器通过注意力机制（权重）+ 深度学习（试错法）获得了信息组织方式的"框架"。这些"框架"就是泛化的基础。"框架"和"框架"之间的映射关系，代表了"矢量"到下一个"矢量"的算法，输入"框架"+不同的细节，就是具体的输入矢量，通过"矢量"到下一个"矢量"的算法，就可以获得输出矢量，这就是知识泛化过程。

但需要注意的是，机器通过"框架"到"框架"的映射，可能产生并不存在的"事实"！ 比如，机器发现很多记者的简介后面，都会有记者的其他文章网页链接，或者附上记者过去获得的奖项。如果机器见到这种信息组织模式很多，那么这种信息组织模式就会成为"框架"到"框架"的映射。所以如果输入信息中包含了类似的框架，但只是记者名字不一样，那么机器都可以通过"框架+细节"，映射到"框架+细节"，从而在输出也产生很多网页链接，或者是奖项。但这些网页链接和奖项也是通过其他"框架+细节"映射到"框架+细节"建立的，它们很可能根本就不存在！

这些问题，大模型难以解决。一种方案是：RLHF 可以用于打碎这种"框架"到"框架"的映射，让机器不要产生对应的文章网页链接或者奖项，但这同时也降低了大模型的能力。另外一种方案是：通过搜索引擎，把和输入相关的信息，加入到用户的输入问题中，从而使得机器获得的输入中包含更多的细节，这样在"框架+细节"映射到"框架+细节"的过程中，产生更加个性化的知识。同样，这是一个治标不治本的解决方案，因为搜索引擎获得的知识本身就不一定正确，而且针对某一个具体问题，搜索引擎能获得的知识很有限。

所以，我们认为 RLHF 是一种解决方案，但它不是最终解决方案。

4 注意力机制+深度学习+强化学习，是走向通用人工智能的正确道路吗？

大模型就可以实现通用人工智能了吗？我们认为答案是否定的。

深度学习就是从大量的样本中，获得一套优化的坐标基底簇。并使用这样的坐标基底簇来表达矢量。所以，把深度学习和注意力机制结合起来后，就能产生类似于人类表达方式的优化的坐标基底簇。这就是 Transformer 能产生智力“涌现”的真正原因。

在 NLP 方面，人类从早期的词袋模型、词向量到 EMLO[21]，直到 Transformer，才真正地实现了注意力机制，并和深度学习无缝地结合起来了，创建了难以置信的奇迹。这就是 LLM。我们注意到，这些技术采用的道路是“先矢量化，建立初步关系；然后通过试错法，来调整坐标基底簇；然后在优选的坐标基底簇下，再次矢量化，获得正确的关系”。这样的机制，导致需要的数据量极大，并且知识是通过训练过程一次成型，难以实时更新的问题[2][3]。

首先，我们注意到，深度学习和人类学习存在两个不同点：（1）它采用的最小信息元素不一样，人类是用自己能感知到的最小局部特征来作为信息空间中的基础元素。而深度学习是用它方便使用的像素、音节或者动作模式作为基础元素。而且这些基础元素实际表现为数据串在时间、空间上的排列。所以深度学习建立的矩阵中最小信息单元（元素），尽管可能和人类的元素相似，但可能并不相同。也许，它可能找到一套更加简洁和高效的元素。

同理，深度学习基于这些最小信息单元所创造的基底坐标簇，同样可能和人类的概念不一致，人类难以理解。但它同样可能建立了一套更加简洁和高效的概念。所以，深度学习确实是一个优雅的解决方案！但它更加适合机器的世界。当人类以人类的标准来评价机器时，我们会就认为机器有时候会犯智障。

其次，上面这个问题，在引入注意力机制后，确实能解决部分问题。通过注意力机制，深度学习获首先关注到最小信息单元之间的关系，并以这些关系为基础来创建基底坐标簇。但由于机器面对的是数据，从数据中获得的最小信息单元，依然可能和人类的最小信息单元不一致，所以机器所创建的“概念”依然可能和人类存在很大差异，这是导致机器不能真正懂得语言的核心问题！

机器的知识，人类很难理解其含义，但机器可以使用，看上去问题不大。但实际上问题很严重。

最核心的问题就是：人类无法模仿机器建立的知识网络形式，给机器预置一些先天知识！这是问题的核心。因为无法给机器预置先天知识，所以不可能模仿机器建立的知识网络形式，给机器预置一些基础需求知识。

目前人工智能最大的缺陷是机器没有自身的需求。因为没有自身的需求，所以不会自发产生目标。没有自发的目标，就不可能有自发的行为。而机器有自发的行为，就是机器给自己编程。能给自己编程的机器，才是真正的智能机器。而需要靠外部编程的机器，始终都是人类智慧驱动的机器。

如何建立机器的需求？首先，机器的需求本身就必须是知识的一部分。因为只有这样，机器才能根据自身知识，和环境互动决策，来满足自身的需求。所以，需求就是知识。

为了实现需求就是知识，我们需要模仿记忆库中最终网络形式，给机器预置一个先天的“最小利弊”内核。通过“利弊内核+小样本学习+不断累积”，最终形成 “带利弊信息的全连接知识网络”。有了“带利弊信息的全连接知识网络”，机器根据自己的知识，自主预测各种决策路径下，可能带来的奖罚值。所以机器就可以按照趋利避害的原则，自己做出决策。

所以机器的所有目标，是机器自己创建的！只有这样，机器才可能在复杂的

环境下，自顶而下，根据具体情况，自主在现场创建子目标，自主决策，自主完成任务！一切预先把任务分解后，再执行的 AGI，它们依然是程序驱动的，无非就是程序的接口可以采用自然语言而已。它们会在真实世界中层出不穷的意外情况下不断碰壁！

有了能自我编程的机器，还必须配套相应的知识，才能真正实现和真实环境互动决策的流程。目前机器和环境互动决策的知识，主要通过强化学习来完成。而强化学习只适用于那些可以大量试错的领域。而面对难以试错的领域，比如照顾病人，比如农业生产，又该如何处理呢？

十多年前，我们专利的几位发明人，讨论应该按照人类学习模式，采用小样本学习来建立人工智能。所以，一开始也是试图走“符号表达”→“常识 + 因果逻辑”→“知识网络”。 尝试几年后，发现这条路的第一步就走不通。因为“符号表达”→“狗”怎么表达？需要把“狗”的所有特征挑选出来。但“狗”可以是一个动物，也可以是一个人！可以是“一种被歌颂的性格”，还可能是“一种被鄙视的性格”....“狗”和“狗腿子”差十万八千里！所以“狗”的本质是“狗”和其他所有事物关系的总和。这是模仿马克思对人的定义来定义“狗”。所以“狗”，必须放到整个知识网络中，通过它和其他所有知识的关系来定义。所以，“符号主义”走不通！因为“狗”不能从其他知识中分割出来！必须建立类似于深度学习的“全连接知识网络”，这是我们的第一个结论。

因为“狗”，必须放到整个知识网络中，通过它和其他所有知识的关系来定义。所以必须要有足够的知识，才能把“狗”这个事说明白。所以，“知识数量必须要足够“，这样才能通过足够的背景知识来理解什么是狗。这是我们的第二个结论。我们回头一看，这不就是大模型干的事吗？“注意力机制 + 深度学习”就是干全连接网络这事，大模型就是干“使用大量知识，来建立全连接知识网络”这事。

那么，为什么我们没有看到满大街走动的机器人？因为只有知识网络还不行！还必须要能够“和环境互动，能连续决策”！而目前的人工智能，和环境互动的决策知识都是靠强化学习来训练。目前 AIXI 算法，本质是理想化的最强“强化学习”算法，它需要的计算量，预测比宇宙的原子都还多，是无法实现的。“alpha go“ 就是采用 AIXI 算法，通过“蒙特卡洛树搜索”对运算量进行修剪，才把下围棋需要的计算量降下来。所以，要想走向通用人工智能，一条可能道路就是：大模型 + AIXI 算法（地球人最强的强化学习算法）。那么，为什么我们没有看到 Google 推出满大街走动的机器人？

这条路的核心障碍是，AIXI 算法需要的两个前提条件[24]：（1），机器需要知道不同决策路径下，它能获得的奖励信息。（2），机器需要遍历搜索所有决策可能性。这两个条件，在游戏里能完美满足。因为游戏里最终输赢就是奖励函数。通过玩一千万次，就能总结出每一步决策的利弊了，这就是决策知识；并且游戏的决策知识搜索空间限制在游戏里，所以机器需要的算力是有上限的。但现实生活中，人生只有一次，不可能无限复盘来获取互动决策经验；而且和游戏不一样，面对一个任务，并没有明确限定的信息搜索范围！所以训练机器打游戏、元宇宙、语言文字图像生成可以试错，但在真实环境里，开车、做饭、照顾孩子、护理病人呢？这些无法试错的领域，目前的人工智能技术道路就解决不了！

5 什么样的道路，才是走向通用人工智能的正确道路？

我们认为，通过真正的“机器学习”实现真正的“通用人工智能”，需要 3

个前提条件：

前提条件 1：足够知识 + 全连接网络，不能有任何外挂！ 任何外挂，和知识网络都无法融为一体，容易偶发智障！

前提条件 2：让机器自己来预测各种决策路径下的奖罚信息！所以机器必须像人一样：自己能预测各种决策路径下的奖罚值，只需要做少量尝试；不是啥事都要去试验一百万次！

前提条件 3：让机器可以直接学习人类所积累的经验！目前主流技术的本质是试错，这叫“机器进化”！人类从单细胞生物那样的智能，进化到今天，用了几亿年！所以机器必须要能够直接学习人类文明史所积累的经验，不能再走“进化”的老路！

我们历时十年，在真正“通用人工智能”需要的 3 个前提条件指引下，提出了一套技术方案[25][26][27][28]。它通过真正的机器学习，实现真正的“通用人工智能”，主要包括：

（一），建立一套人类可以理解的全连接知识网络。而建立的方法，就是通过注意力机制和记忆和遗忘机制来实现的。记忆和遗忘机制主要实现统计关联，而信息的注意力机制则是在统计关联的基础上，通过链式联想激活过程，通过激活值多路径累积，通过激活值随时间消退来实现的。

（二）因为我们的全连接知识网络是人类可以理解的网络组织形式（事实上，它就是一个数据库）。所以，我们可以模仿记忆库中最终网络形式，给机器预置一个先天的“最小利弊”内核。通过“利弊内核+小样本学习+不断累积”，最终形成 “带利弊信息的全连接知识网络”。有了“带利弊信息的全连接知识网络”，机器根据自己的知识，自主预测各种决策路径下，可能带来的奖罚值。所以机器就可以按照趋利避害的原则，自己创建目标，自己做出决策。只有这样，机器才可能面对复杂的环境下，自顶而下，根据具体情况，自主创建目标，自主决策，自主完成任务!

（三）机器面对的环境千差万别，机器面对的任务也是千差万别，所以，机器不可能在任意环境下，通过训练来获得任意任务的相关互动决策知识！这是一个不可能完成的任务！

所以，我们必须转变思路。从人类的学习获得启示。事实上，人类面对任务时，所有的决策都是围绕趋利避害这个核心来决策的。所以才会有回避，拒绝，寻找更多的帮助等行为。这些行为本质上是人类创建的新行为。人类不是直接去处理任务，而是把任何任务转换为“如何满足自身需求”的问题。

同理，机器也必须这样。让机器在日常生活中，不断学习“如何满足自身需求”的过程。因此，针对任意任务，机器都是按照自身需求，把它转换为“如何满足自身需求”这个任务。而在这个任务上，它有大量可以泛化的经验，因为它的所有学习过程，都是围绕这个任务进行的。所以，我们提出的解决方案是：机器的学习过程，不能以任务为导向，而是应该以机器自身需求为导向。如果机器拥有自身的需求，并且拥有如何满足自身需求的相关知识，那么机器就可以创建出新的行为，来满足自身的需求。也就是说，机器实现了自己给自己“编程”，而且需要编程的任务只有一个“如何满足自身需求”，机器的所有知识也都是围绕“如何满足自身需求”来建立的。所以，我们的机器，才能完成任意任务。它是在完成“如何满足自身需求”的过程中，也对完成了人类给予的具体任务。完成的结果有可能是“完成”、“拒绝”或者“进一步寻找更多的信息来做评估”。

如果我们给机器预置了多种对人类正面反馈的先天需求，那么机器就是安全

的，而且它也会主动一步步分解任务和完成任务。

所以，机器自始自终，都在学习完成一件任务“如何满足自身需求”，它也自始自终，都在处理一件任务“如何满足自身需求”。而完成具体的任务，这是处理“如何满足自身需求”任务途中的副产品。

所以，建立真正的通用人工智能，一条可行的道路就是：模仿机器的知识组织形式，预置机器的需求。然后，机器创建的知识中，需要包含需求相关的知识。

什么是知识？就是常见特征在时间、空间中的排列方式！什么是需求相关的知识？就是常见特征在时间、空间中的排列方式，这种排列方式中包含了机器的需求信息。常见特征在时间、空间中的排列方式，如果包含了机器的需求，就是主观常识。也就是“世界”和“我”的关系。常见特征在时间、空间中的排列方式，如果不包含了机器的需求，就是客观常识。也就是“万物之间”的关系。所以，常见特征在时间、空间中的排列方式，就是常识。要实现真正的通用人工智能，最核心的就是要实现“常识”！有了常识，机器才会根据自身需求，主动去解决任务，主动去创建流程。也就是说，机器给自己编程！这才是真正的智慧！而目前大模型+万物 APP 的方式，依然没有脱离人类编程的方式！

（四）机器需要把知识网络 + 机器需求 + 价值评估融为同一个网络。机器在日常生活中，就是不断学习“如何满足自身需求”的过程。因此，针对这个问题，它有大量可以泛化的经验。如果我们给机器预置了多种对人类正面反馈的先天需求，那么机器就是安全的，而且它也会主动一步步分解任务和完成任务。

先天需求里面，需要包含机器自身的运作需求，这样机器就会去维护自身的运转。也需要包含人类给机器预置的需求，比如机器渴望获得人类的正面反馈，这和人类小孩一样。通过这样的方式，我们就可以和机器互动，从小训练，让机器的价值观和人类对齐。建立起机器的本能需求，也建立起机器的高阶价值观需求，比如“道德观”，“遵守法律”等。同时，我们还可以预置少量先天知识，这些先天知识主要是用于不可试错的领域，比如最小的“悬崖躲避”知识。这样，我们就创造出一个孩子，它有需求，它自私，它有少量和生存相关的本能知识。但它渴望人类的认可。它有一套和人类沟通的先天语言（比如只需要认识点头或者摇头的先天知识就可以了）。然后，基于和人类沟通的先天语言，人类可以慢慢和机器建立更加复杂的沟通方式，比如语言。然后，机器通过在实际环境中学习，既可以通过自我总结获得知识，也可以通过语言学习直接获得人类已有的知识，还可能利用它无与伦比的强大能力发现人类没有发现的常见特征排列方式，尽管这些特征在时间和空间上的排列模式，对人类并不明显，但机器可以通过统计发现，并且可以模仿人类采用符号来表达常见排列的方式，采用符号来表达新发现的常见特征排列方式。这就是机器创建的新知识。

这是一个迭代过程。机器给自己编程，让自己发现更多的知识，它将会发展为超级智能！

（五）机器需要通过小样本、累积学习。这样才可能实现知识的实时更新。

在现实生活中，大量的任务需要通过机器和环境互动来一步步完成。所以，环境的任何反馈，都必须立即成为机器下一步决策的依据。并且这种决策知识，需要立即更新到机器的知识库中。否则，机器下一次会犯同样的错误。而目前人工智能，采用大数据样本，知识主要通过一次训练完成。即使是针对任务的微调，也无法实现实时更新。所以，真正的学习道路，应该是小样本、累积学习。这样的学习方式，和人类更相似，才可能实现实时更新。

5 人工智能的演进方向初探

我们认为，人工智能的发展可以近似分为不同的阶段：（1）在实现真正的注意力之前的阶段，可以认为是“特征探索”阶段。深度学习之前，主要集中在“人工探索”阶段。人工探索有可以是“专家系统”、“知识百科”、“概率统计”等方式。在深度学习之后，集中在“机器探索”阶段，让机器从大样本中自己“探索特征”。（2）在实现了真正的注意力（Transformer）之后，因为机器的“知识”和人类“知识”初步对齐后，可以认为机器实现了“知识泛化”。面对人类的任务，机器可以通过“知识泛化”表现出一定的智能。在未来，我们认为人工智能需要发展到下一个阶段：“自主互动”阶段。“自主”意味着机器不再是沉默的“机器”，它能够自发地产生行为（这等同于给自己编程），机器会自我探索知识（比如主动和环境互动，获得知识）。“互动”意味着机器可以和环境实时互动，实时更新自己的知识，并能进行连续决策，在陌生环境下完成复杂的任务。

而实现“自主互动”的核心是机器需要有自身的需求。机器的需求必须是机器知识的一部分，这样机器才可能利用自身的知识，来创建行为，从而满足自身需求。

而实现机器需求的核心是首先要创建人类可以理解的知识网络。只有这样，人类才可以模仿知识网络的形式，预置机器的需求。然后让机器围绕自身的需求来学习，从而建立起所有信息和需求之间的连接关系。所以机器的知识都是和需求相关的。这样，机器才可以把任何具体任务转换为“如何满足自身需求”这样的单一任务。而机器的所有探索和学习过程，也都是围绕“如何满足自身需求”这个任务展开的。所以机器面对“如何满足自身需求”时，有大量可以泛化的经验。只有这样，机器才可能处理各种难以试错的任务。事实上，我们认为，人类也是使用类似的方法来获得知识和处理问题的。

人工智能面向“自身需求”，而不是面向“外部任务”是一种范式的转变。我们认为这种范式转变是必要的。因为外部任务千差万别，大量的任务必须和实际环境互动。它们难以试错，就难以获得大数据样本，就难以通过强化学习来获得决策知识。另外，让机器针对每种类型的任务进行训练，本身也是一个不可能完成的任务。

通用人工智能是人工智能的初心，也是人工智能的桂冠。我们提出了一套实现通用人工智能的技术方案。在参考文献[25][26][27][28]中，我们提出了实现这条道路的详尽技术细节，它有可能是一条引导人类走向通用人工智能的可行道路。

在这套方案中，机器的需求是由人类预置的，是可以有多种需求的，所以机器产生的目标也是多目标的。目前人工智能是单一目标，从性格来讲，可以认为它就是“为了目标，不择手段”的人工智能。因为这样的人工智能，它只追求单一目标，而不会去考虑目标之外的任何东西。所以，这是非常危险的人工智能！而在我们的方案中，多目标也包括对齐人类的价值观，包括“道德”、“法律”、“认可”等需求，所以我们的方案中，机器会综合考虑“道德”、“法律”、“认可”等需求，所以我们的方案，是一条解决人工智能安全性的可行道路。

参考文献：

[1] A Generalist Agent, Scott Reed, Konrad Zona , Emilio Parisotto , Sergio Gomez Colmenarejo , Alexander Novikov, Gabriel Barth-Maron, Mai Giménez, Yury Sulsky,

Jackie Kay, Jost Tobias Springenberg, Tom Eccles, Jake Bruce, Ali Razavi, Ashley Edwards, Nicolas Heess, Yutian Chen, Raia Hadsell, Oriol Vinyals, Mahyar Bordbar and Nando de Freitas, https://openreview.net/forum?id=1ikK0kHjvj

[2] OpenAI. ChatGPT: Optimizing Language Models for Dialogue. https://openai.com/blog/chatgpt/, 2023

[3] Ouyang L, Wu J, Jiang X, et al. Training language models to follow nstructions with human feedback[J]. arXiv:2203.02155, 2022

[4]Vinyals O, Ewalds T, Bartunov S, et al. Starcraft ii: A new challenge for reinforcement learning[J]. arXiv:1708.04782, 2017.

[5] A Survey of Learning-based Automated Program Repair, Antonio Mastropaolo, Luca Pascarella, Emanuela Guglielmi, Matteo Ciniselli, Simone Scalabrino, Rocco Oliveto, Gabriele Bavota, arXiv:2302.00438

[6] A Survey of Learning-based Automated Program Repair,Quanjun Zhang, Chunrong Fang, Yuxiang Ma, Weisong Sun, Zhenyu Chen,arXiv:2301.03270

[7] Introducing ChatGPT, https://openai.com/blog/chatgpt/

[8] Prompting GPT-3 To Be Reliable,Chenglei Si, Zhe Gan, Zhengyuan Yang, Shuohang Wang, Jianfeng Wang, Jordan Boyd-Graber, Lijuan Wang, arXiv:2210.09150

[9] An experimental open-source attempt to make GPT-4 fully autonomous, https://GitHub.com/Significant-Gravitas/Auto-GPT

[10] https://jina.ai/news/auto-gpt-unmasked-hype-hard-truths-production-pitfalls/

[11] https://github.com/Torantulino/Auto-GPT?ref=jina-ai-gmbh.ghost.io

[12] Auto-GPT - The next evolution of data driven Chat AI, https: // auto -gpt.ai

[13] 文心一言：https://mp.weixin.qq.com/s/0-8X9FPouteKzNiK6DPaiA

[14] Learned in translation: Contextualized word vectors. In Advances in Neural Information Processing Systems

[15] Attention Is All You Need，Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin，arXiv:1706.03762

[16] Jianpeng Cheng, Li Dong, and Mirella Lapata. Long short-term memory-networks for machine reading. arXiv preprint arXiv:1601.06733, 2016.

[17] Zhouhan Lin, Minwei Feng, Cicero Nogueira dos Santos, Mo Yu, Bing Xiang, Bowen Zhou, and Yoshua Bengio, A structured self-attentive sentence embedding. arXiv:1703.03130, 2017.

[18] Ankur Parikh, Oscar Täckström, Dipanjan Das, and Jakob Uszkoreit. A decomposable attention model. In Empirical Methods in Natural Language Processing, 2016.

[19] Romain Paulus, Caiming Xiong, and Richard Socher. A deep reinforced model for abstractive summarization. arXiv preprint arXiv:1705.04304, 2017.

[20] Tamkin A, Brundage M, Clark J, et al. Understanding the capabilities, limitations, and societal impact of large language models[J]. arXiv:2102.02503,

2021.
[21] BERMo: What can BERT learn from ELMo?Sangamesh Kodge, Kaushik Roy,arXiv:2107.03508
[22] OpenAI. ChatGPT: Optimizing Language Models for Dialogue.https://openai.com/blog/chatgpt/, 2023
[23] Ouyang L, Wu J, Jiang X, et al. Training language models to follow instructions with human feedback[J]. arXiv:2203.02155, 2022
[24] Principles of Solomonoff Induction and AIXI,Peter Sunehag, Marcus Hutter,arXiv:1111.6117
[25] 一种实现通用人工智能的方法， 陈永聪，曾婷，陈星月，中国专利CN111553467B
[26] 一种模仿人类智能的机器智能实现方法陈永聪，曾婷，陈星月，中国专利CN111563575B
[27] 一种实现类人通用人工智能机器的方法，陈永聪，曾婷，陈星月，中国专利CN112215346B
[28] ESTABLISHMENT OF GENERAL-PURPOSE ARTIFICIAL INTELLIGENCE SYSTEM，陈永聪，张俊，曾婷，陈星月，美国专利US20220121961A1